**Data rensning**

Data er hentet fra ICES’ database of trawl surveys (DATRAS) (link). For at dække hele Danmark, er to surveys brugt: North Sea Internaional Bottom Trawl Survey(NS-IBTS) og Baltic International Trawl Survey (BITS). (indsæt billede af hvor surveys foregår).

7 polygoner/områder er laver i Q GIS og data er klippet til hvert område. Dvs. der nu er 7 DATRAS datasæt: et over hvert område. Hvert område er valgt ud fra tilgængeligheden af data; der skal være nok hauls der går langt nok ”tilbage i tiden”. (evt indsæt screenshot??)

Derudover lægges en shape fil fra GEUS og … om havbundssedient ud og klippes også til data samt … for at tilføje en kolonne i attributtabellen om sedimenttype. Hvorfor?

Shape Filerne af polygonerne kan hentes i github.

I github hedder data-filerne: (kan evt stå en bjælke ude i siden hvor der står hvad det er) dat\_øresund etc og er csv filer

(note: valgene til rensning af datasæt er taget efter mange timers deskriptiv statistik (arter, tid, kvartal, etc. )

Nu laves der data rensning i R over arterne: brisling, sild, skrubbe, rødspætte, torsk, hvilling. De her arter er valgt fordi der er tilstrækkeligt med data på dem og de er vigtige kommercielle arter…

Hver art inddeles i tre størrelser ud fra biologiske kriterier…

For hvert område\_art\_størrelse renses der således:

* Kun kvartal 1 og 4
* Kun >år 2000
* Der laves en ny kolonne med et ”unikt” hauls ID, da der ellers tælles fra 1 ved hvert nyt år og land og skib. For at lave unik haul id bruges ”Year”, ”HaulNo”, ”Quarter”, ”Survey” ”country” “Ship” og “Gear”.
* Filtrer til den specifikke art man er ved at undersøge
* Sørger for at alle længdeklasser er repræsenteret ved at tilføje 0 i de klasser der ikke er fanget i ahulet
* Laver nye data sæt som kun indeholder den valgte størrelse af arten
* Tilføj ny kolonne med den summerede CPUE af arten i område i størrelsesklassen.

**GAM**

Der er en nu 6 csv filer der hedder: dat\_art etc

De bliver uploaded i ”OmrådeStørrelse.R” filerne og modellerne for hvert art, størrelse og område kan nu laves

Modellerne består ser ud som følger:

Indsæt billede fra R

Beskrivelse af parametre:

* cpue:
* period:
* s(depth)
* s(Gear)
* Quarter
* SedimentDK

For at få den bedste, mindste model bruges scriptet ”funcs” som ….. tOBIAS?

Med modellerne udregnes nu en predicted CPUE og en predicted sd oog gemmes som ”newdat” rds, så de kan bruges i shiny appen

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Kattegat N | Kattegat, Læsø | Nord for øresund | Øresund | Storebælt | Mecklenburg bugten | Sydvest for Bornholm |
| Adult sprat | Juvenile fliunder | Whinting A,  Flounder M  Flounder J | Flounder J, plaice J | Cod A, sprat A, whinting A, flounder J | Plaice J, sprat J, flounder J, cod A, sprat A, whinting A | Cod A, sprat A, whiting A |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Kattegat N | Kattegat M | NTZ | Øresund | Storebælt | Mecklenburg Bugten | Østersøen |
| 154 | 225 | 77 | 143 | 231 | 372 | 260 |

**Plots**

Prikken er udregnet som et forhold mellem middelværdierne i hver periode for cpuen. Whiskers er udregnet med sds.. indsæt funktionen. Farverne er givet alt efter om der overlappes ved 0.